AN EFFICIENT REAL-TIME DATA COLLECTION MODEL FOR MULTIVARIATE SENSORS IN INTERNET OF THINGS (IOT) APPLICATIONS

NAYEF ABDULWAHAB MOHAMMED ALDUAIS

A thesis submitted in
fulfillment of requirement for the award of
Doctor of Philosophy in Electrical Engineering

Faculty of Electrical and Electronic Engineering
Universiti Tun Hussein Onn Malaysia

MAY 2019

For my beloved mother, father, wife, sons, brothers, sisters, family

ACKNOWLEDGEMENT

First of all, praise and thanksgiving to Allah for the blessing of mind and health and “Tofiqime”to complete this thesis. In spite of the fact that my name is printed on the cover of this thesis, the word” I” does not appear within its chapters. I do this to pay tribute to the myriad contributions of my supervisors and mentors, and the support of my family and friends. To my supervisors Assoc. Prof. Dr. Jiwa bin Abdullah and Dr. Ansar Jamil, thank you for your constant guidance and endless patience towards the completion of this thesis. Without your insights and encouragement, this thesis would not have been completed. To my parents Abdullwahab and Fauzia: I owe you for your love, affection and sacrifice. Thank you for giving me the freedom and opportunity to pursue my studies, and forgive me for not being there, when you needed me the most. To my wife Salaw, thank you for believing in me, and for your patience during so many late nights, many of them unanticipated. To my sons Nawaf and Muntaser: Every morning when you wake me up, I feel joy and inspiration. Thank you for things which you do not even know you have done for me -I can even write another thesis. To my brothers and sisters: You are always happy with my every success, thanks a lot for your moral supports. To my friend Adel Yahya Ashype thanks for everything. To UTHM University and the staffs of Faculty of Electrical and Electronic Engineering: Thanks for providing me with an excellent research environment and the necessary resources to undertake this research. To Yemen government and Hodeidah University: Thank you for support me. Last but not least, I would like to thank a person who contributes to complete my final thesis report directly or indirectly. I would like to acknowledge him/her helps, which was necessary to complete this.

ABSTRACT

In the applications of the Internet of Things (IoT), sensor board depends on a battery that has a limited lifetime to function. Furthermore, the IoT sensor board with multivariate sensors influences the battery lifetime since there is additional data transmissions that must be supported by the board causing it to drain the battery much faster than the sensor board with one sensor. The main aim of this thesis is to increase the battery life of the IoT sensor node. To do so, a number of proposals are presented. First, an updating data strategy denoted as an efficient data collection and dissemination (EDCD) is proposed. EDCD aims to save the energy consumption of the IoT sensor board with multiple sensors by means of reducing the number of transmission packets, if no significant change is reported by the payload sensing block; second is proposed a validity of the measuring sensor reading at node level (VSNL) algorithm. VSNL aims to avoid transmitting any incorrect data, which will help in saving the energy consumption; third, an adaptive threshold and new metric for multivariate data reduction models such as principal component analysis – based (PCA-B) and multiple linear regression – based (MLR-B) have been proposed. In addition, proposed a payload data reduction algorithm (APRS). APRS aims to reduce the transmitted packet size for each sensed payload, which that will help in saving the energy of the IoT sensor board. This work provides an extensive analysis for the design and performance evaluation of real-time data collection model for multivariate sensors in IoT applications. Finally, an efficient real-time data collection model for multivariate sensors in IoT applications (RDCM). RDCM integrated EDCD, VSNL, PCA-B/MLR-B and APRS and the ability to prolong sensor board battery lifetime, which that satisfied by reducing number of transmissions and payload packet size, and also increase the accuracy of data validation. Performance of the proposed algorithms was evaluated through simulation by utilising various real-time datasets. The average of the total percentage of energy saved by applied RDCM to real-time data sets injected with various percentage of errors for all nodes is 98%.

ABSTRAK

Dalam aplikasi Internet of Things (IoT), papan sensor bergantung kepada bateri yang mempunyai jangka hayat yang terhad untuk berfungsi. Selain itu, papan sensor IoT dengan sensor pelbagai variat mempengaruhi jangka hayat bateri memandangkan terdapatnya penghantaran data tambahan yang perlu disokong oleh papan tersebut sehingga ia menyebabkan pengurangan bateri berlaku dengan lebih cepat daripada papan sensor dengan satu sensor sahaja. Untuk mencapai matlamat utama tesis ini; pertama, strategi pengemaskinian data yang dikenali sebagai Algoritma Kutipan Data dan Agihan (EDCD) telah dicadangkan. EDCD bertujuan untuk menjimatkan penggunaan tenaga oleh papan sensor IoT dengan pelbagai sensor dengan cara mengurangkan bilangan penghantaran paket, jika tiada perubahan penting dilaporkan oleh blok sensor muatan; kedua, semak kesahihan membaca bacaan pengukur di peringkat nod (VSNL) telah dicadangkan. VSNL mensasarkan untuk mengelakkan penghantaran data yang salah, yang akan membantu dalam penjimatan penggunaan tenaga serta meningkatkan ketepatan sistem; ketiga, satu nilai ambang mudah suai dan matrik baharu bagi model pengurangan data pelbagai sumber yang berasaskan kepada penganalisis koponen utama (PCA) dan yang berasaskan kepada linear regressi berbilang adalah di cadangkan. Seterusnya juga dicadangkan satu algoritma pengurangan data muatan di kenali sebagai APRS. APRS bertujuan untuk mengurangkan saiz paket yang akan dihantar bagi setiap muatan data sensor yang boleh menjimatkan tenaga papan sensor IoT. Kerja ini menyediakan satu analisis yang terperinci untuk reka bentuk dan penilaian prestasi model pengumpulan pelbagai variasi data masa nyata dengan pelbagai variasi sensor untuk aplikasi IoT. Akhirnya, mencadangkan Model Pengumpulan Data Nyata untuk papan sensor IoT dengan pelbagai variasi sensor yang dipanggil RDCM. RDCM bertujuan untuk memanjangkan jangka hayat bateri bagi papan sensor, yang selari dengan pengurangan bilangan penghantaran dan saiz paket muatan, dan juga meningkatkan pengesahan data

yang tepat. Sumbangan RDCM berdasarkan skim-skim tersebut (EDCD, VSNL, PCA-B / MLR-B dan APRS). Prestasi algoritma yang dicadangkan dinilai melalui simulasi dengan menggunakan pelbagai dataset masa nyata. Purata peratusan penjimatan tenaga oleh RDCM yang diaplikasikan untuk set data masa nyata yang disuntik dengan pelbagai peratusan kesilapan untuk semua nod adalah 98%.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$D[t]$	-	the representation variable real-time sensed (reduced sensed data)
D_S	-	Node Data size (D_S) is all the measured values read by the all sensors S_i , i=1,2...,n
E	-	the total energy consumption during a specific period.
ES	-	Abbreviation to Exactly same (ES), the value of absolute change between ($S_{V_{(t-1)}}$) and ($S_{V_{(t)}}$) is "zero".
$E_{R,}E_T$	-	the energy consumption for transmission of data in the reduction phase and retraining phase, respectively.
E_{RM}	-	the energy consumption per sample in the Reduction mode (RM)
E_{Byte}, E_{bit}	-	the energy consumption per Byte and bits, respectively.
E_{N-RM}	-	the energy consumption per sample in the Non-Reduction Mode (N-RM)
E_{total}	-	the total energy consumption during a specific period.
F_S	-	the binary parameter to send data / not send data
R_f	-	The relative difference
K	-	the total number of transmissions samples
K_1 , K_2 K_3	-	the number of transmit samples during data collection in N-RM, RM, and RTM, respectively.
L	-	the total number of bits required to represent relative difference $\pm RD_i$
$L1$	-	the total number of bits for D[t] at the fusion center
m	-	the required bits to represent $\|RD_i\|$
n	-	the number of sensors on the same IoT sensor board
n_R, n_T	-	the number of message transmissions in the reduction phase, the number of UFM
Nd	-	Number of nodes

P_{Lenght}	-	the length (bits) of payload per sample
PC	-	the number of Principal Components (PC) for PCA
$r\%$	-	the data reduction ratio
$RT_n\%$	-	the average percentage of power saved in transmission phase
$RD_{1\times n}$	-	the model data reduction ratio
$S(t)$	-	the current measured value by the sensor
$S(t-1)$	-	the last measured value transmitted by the sensor.
S_i	-	Array of sensors and *i-th* Refer to the index of sensor
S_{s_i}	-	State of the sensor $(S_{S_i}) \in [1,0]$ is a one-bit value to describe the correlation between the current measuring (S_t) and the previous transmitted measuring value(S_{t-1}).
S_s	-	A row data for sensors state value
S_d, S_R	-	the size of the original sensed data and of model reference parameters
$S_{1\times n}\,[t]$	-	the new Real-time sensed data vector at the time t
$S_{1\times n}\,[t-1]$	-	the last sensed data transmitted vector to the fusion center at t-1
$Sb_{1\times n}$	-	the signs vector in order to manage the negative and none-negative RD
$\hat{S}_{1\times n}[t]$	-	the predictions sensed data at Sensor board / Fusion
t	-	refer to the current time
$t-1$	-	refer to the last time transmitted to the sink
β	-	the allowed amount of difference (select by the Sink)

LIST OF ABBREVIATIONS

ABSD	-	An Adaptive Buffer Sensed Data
ADAT	-	A technique to Reduce Aggregated Data
AM	-	Arithmetic Mean
ANN	-	Artificial Neural Networks
APCADR	-	Adaptive PCADR
APRS	-	Adaptive Real-Time Payload Data Reduction Scheme
AR	-	Autoregressive
ARIMA	-	Autoregressive Integrated Moving Average
AT	-	Ambient Temperature
CCIPCA	-	Candid Covariance-Free Incremental PCA
CH	-	Cluster Head
CMF	-	Confidence Measure Factor
CoPeST	-	Collective forecast exploiting temporal-spatial correlation
CSE	-	Constant Error
cor	-	Correlation
DVA	-	Data validation algorithm
ED	-	Euclidean distance
EDCD	-	An Efficient Data Collection For IoT /WSN Applications
ED-HD	-	Error Detection based Historic Data
ED-NN	-	Error Detection based Nearest Neighbors
ED-RT	-	Error Detection in Real-Time
ES	-	Exponential Smoothing
EV	-	Event
FD	-	False Detection
GM	-	Geometric mean
GUI	-	Graphic user interface
HLMS	-	Hierarchical Least-Mean-Square
HM	-	Harmonic mean
IBRL	-	Intel Research Lab dataset
IoT	-	Internet of Things

KPCA	-	kernel principal component analysis
LMS	-	Least Mean Square
LUCE	-	Lausanne Urban Canopy Experiment
LZW	-	Lempel–Ziv–Welch
MA	-	Moving Average
MCU	-	microcontroller
MD	-	Median
MLP	-	multi-layer perceptron
MLR / LR	-	Multiple / simple linear Regression
MLR-B	-	Multiple linear Regression –based Model
MSD	-	mean-square derivation
NC	-	Normal change
OCPCC	-	One Class Principal Component Classifier
OCSVM	-	one class quarter sphere support vector machine
OTE	-	Outiler Error
PCA	-	Principal Component Analysis
PCA-B	-	Principal Component Analysis – based Model
PCADR	-	PCA Data Reduction
RD	-	Relative Difference
RDCM	-	An Efficient Real-Time Data Collection Model For Multivariate Sensors in IoT Applications
RE	-	Range Error
RH	-	Relative Humidity
RMSE	-	Root Mean Square Error
RTD	-	real-time data
SD	-	Standard deviation
SMG	-	Sensor Measure’s Gradient
ST	-	Surface Temperature
TSF	-	Time Series Forecasting
UFM	-	Updating Frequency Metric
VSNL	-	Validity of the measuring sensor reading at node level
WSN	-	Wireless sensor networks

LIST OF APPENDICES

REFERENCES

1. Atzori, L., Iera, A., & Morabito, G. (2010). “The internet of things:A survey”. *Computer networks*, 54(15): 2787-2805.
2. A. Zanella, N. Bui, A. Castellani, L. Vangelista and M. Zorzi. (2014). "Internet of Things for Smart Cities," *in IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22-32.
3. The Four Internet of Things Connectivity Models Explained", Channel Futures, 2018. [Online]. Available: http://www.channelfutures.com/best-practices/four-internet-things-connectivity-models-explained. [Accessed: 01- Mar- 2018].
4. Mukhopadhyay, S. C. (2015). “Wearable sensors for human activity monitoring: A review”. *IEEE Sensors Journal*, , 15(3), 1321-1330.
5. Rodgers, M. M., Pai, V. M., & Conroy, R. S. (2015). “Recent advances in wearable sensors for health monitoring”. *IEEE Sensors Journal*, 15(6), 3119-3126.
6. Ghayvat, H., Liu, J., Mukhopadhyay, S. C., & Gui, X. (2015). “Wellness Sensor Networks: A Proposal and Implementation for Smart Home for Assisted Living”. *IEEE Sensors Journal*, 15(12), 7341-7348.
7. Gope, P., & Hwang, T. (2016). “BSN-Care: A secure IoT-based modern healthcare system using body sensor network”. *IEEE Sensors Journal*, 16(5): 1368-1376.
8. Kumar, A., & Hancke, G. P. (2015). “A Zigbee-based animal health monitoring system”. *IEEE Sensors Journal*, 15(1): 610-617.
9. Kelly, S. D. T., Suryadevara, N. K., & Mukhopadhyay, S. C. (2013). “Towards the implementation of IoT for environmental condition monitoring in homes”. *IEEE Sensors Journal*, 13(10): 3846-3853.
10. Girish, S. V., Prakash, R., & Ganesh, A. B. (2016). “Real-Time Remote Monitoring of Indoor Air Quality Using Internet of Things (IoT) and GSM

Connectivity". *Artificial Intelligence and Evolutionary Computations in Engineering Systems* ,pp. 527-533.

11. Shi, W., Cao, J., Zhang, Q., Li, Y. and Xu, L. (2016)." Edge computing: Vision and challenges". *IEEE Internet of Things Journal*, 3(5), pp.637-646.
12. Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012, August). "Fog computing and its role in the internet of things". *Proceedings of the first edition of the MCC workshop on Mobile cloud computing* (pp. 13-16). ACM.
13. Satyanarayanan, M., Bahl, V., Caceres, R., & Davies, N. (2009). "The case for vm-based cloudlets in mobile computing". *IEEE pervasive Computing*,(4): 14-23.
14. A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. (2008). "The cost of a cloud: Research problems in data center networks," *ACM SIGCOMM Comput. Commun. Rev*., vol. 39(1): 68–73.
15. Cuervo, E., Balasubramanian, A., Cho, D. K., Wolman, A., Saroiu, S., Chandra, R., & Bahl, P. (2010). "MAUI: making smartphones last longer with code offload". *In Proceedings of the 8th international conference on Mobile systems, applications, and services* (pp. 49-62). ACM.
16. Alduais N.A.M, Jamil.A and Abdullah.J. (2015). "Performance Evaluation Of Different Logical Topologies And Their Respective Protocols For Wireless Sensor Networks". *ARPN J. Eng. Appl. Sci*. 10 8625–34.
17. Zhang, Y., Meratnia, N. and Havinga, P.J. (2010). "Outlier detection techniques for wireless sensor networks: A survey". *IEEE Communications Surveys and Tutorials*, 12(2): pp.159-170.
18. Rassam, M.A., Zainal, A. and Maarof, M.A.(2013). "Advancements of data anomaly detection research in wireless sensor networks: a survey and open issues". *Sensors*, 13(8), pp.10087-10122.
19. Jawad, H. M., Nordin, R., Gharghan, S. K., Jawad, A. M., & Ismail, M. (2017). "Energy-Efficient wireless sensor networks for precision agriculture: A review". *Sensors*, 17(8):1781.
20. Gharghan, S. K., Nordin, R., & Ismail, M. (2014). "Energy-efficient ZigBee-based wireless sensor network for track bicycle performance monitoring". *Sensors*, 14(8):15573-15592.
21. Babaie, S., Khadem-zadeh, A., & Badie, K. (2012). "Distributed Fault Detection Method and Diagnosis of Fault Type in Clustered Wireless Sensor

Networks”. *Life Science Journal*, 9(4).

22. Ayadi, A., Ghorbel, O., Obeid, A. M., & Abid, M. (2017). “Outlier detection approaches for wireless sensor networks: A survey”. *Computer Networks*, 129, 319-333.

23. Sebestyen, G., Hangan, A., Czako, Z., & Kovacs, G. (2018). “A taxonomy and platform for anomaly detection”. *2018 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)* : pp. 1-6.

24. Bispo, K. A., Rosa, N. S., & Cunha, P. R. (2012). “A semantic solution for saving energy in wireless sensor networks”. *IEEE Symposium on Computers and Communications (ISCC)* pp. 000492-000499.

25. Liu, L., Men, S., Liu, M., & Zhou, B. (2014). “An energy saving solution for wireless communication equipment”. *Telecommunications Energy Conference (INTELEC)*. pp. 1-3.

26. Rooshenas, A., Rabiee, H. R., Movaghar, A., &Naderi, M. Y.(2010). “Reducing the data transmission in wireless sensor networks using the principal component analysis”. *Sixth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*.pp. 133-138

27. Lawrence, E., Navarro, K. F., Hoang, D., & Lim, Y. Y. (2009). “Data collection, correlation and dissemination of medical sensor information in a WSN”. *Fifth International Conference on Networking and Services*. pp. 402-408.

28. W.IoT, Definitionfrom WhatIs.com.”[Online].Available: https://internetofthingsagenda.techtarget.com/definition/IoT-device. [Accessed: 01-March-2019].

29. Čolaković, A., & Hadžialić, M. (2018).“Internet of Things (IoT): A review of enabling technologies, challenges, and open research issues”. *Computer Networks*. vol. 144 pp. 17-39.

30. C. Prototyping, C. project, "IoT Hardware Guide | 2017 Prototyping Boards & Development Kit Options", Postscapes.com, 2017. [Online]. Available: http://www.postscapes.com/internet-of-things-hardware/. [Accessed: 02-Jan- 2017].

31. R. Blog , "Without Sensors, There is No Internet of Things (IoT) - Rogers Corporation Blog", Blog.rogerscorp.com, 2017. [Online]. Available: http://blog.rogerscorp.com/2016/07/31/without-sensors-there-is-no-internet-

of-things-iot/. [Accessed: 02- Jan- 2017].

32. Gardašević, G., Veletić, M., Maletić, N., Vasiljević, D., Radusinović, I., Tomović, S., & Radonjić, M. (2017). "The IoT architectural framework, design issues and application domains". *Wireless personal communications*, 92(1), 127-148.

33. Sheng Z, Mahapatra C, Zhu C, Leung VC.(2015). "Recent advances in industrial wireless sensor networks toward efficient management in IoT". *IEEE access*.;3:622-37.

34. T. Libelium , "Top 50 Internet of Things Applications - Ranking | Libelium", Libelium.com, 2018. [Online]. Available: http://www.libelium.com/resources/top_50_iot_sensor_applications_ranking/. [Accessed: 07- Mar- 2018].

35. Akpakwu, G. A., Silva, B. J., Hancke, G. P., & Abu-Mahfouz, A. M. (2018). A survey on 5G networks for the Internet of Things: Communication technologies and challenges. *IEEE Access*, 6, 3619-3647.

36. W. Libelium ,"Waspmote", Libelium. (2019) .ModulesWaspmote Wireless Interfaces Overview - Complete List of Modules |. http://www.libelium.com/products/waspmote/interfaces/

37. S. Technology , Sigfox Technology Overview | Sigfox. (2019). Retrieved from https://www.sigfox.com/en/sigfox-iot-technology-overview

38. L. Alliance ," LoRa Alliance, "LoRaWAN Specification v1.0," Technical Report, Jan. 2015.

39. Margelis, G., Piechocki, R., Kaleshi, D. Thomas, P. (2015). "Low throughput networks for the IoT: Lessons learned from industrial implementations". *IEEE 2nd world forum on internet of things (WF-IoT)* .pp. 181-186.

40. I..Bluetooth, "IEEE 802.15.1 Bluetooth.".(2019). Retrieved from http://www.ieee802.org/15/pub/TG1.html.

41. BSIG, "Bluetooth Special Interest Group," [Online]. Availability: https://www.bluetooth.org. [Accessed: 08- Mar- 2019].

42. IEEE Computer Society LAN/MAN Standards Committee. (2007). IEEE standard for information technology-telecommunications and information exchange between systems-Local and metropolitan area networks-specific requirements part 11: *Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. IEEE Std 802.11*.

43. RPMA, Technology for the Internet of Things.(2019). Retrieved from https://theinternetofthings.report/Resources/Whitepapers/4cbc5e5e-6ef8-4455-b8cd-f6e3888624cb_RPMA%20Technology.pdf

44. S. Instruments ,"Sensor Terminology - National Instruments", Ni.com, 2017. [Online]. Available: http://www.ni.com/white-paper/14860/en/. [Accessed: 03- Jan- 2017].

45. G.Libelium,Gases Board Technical Guide | Libelium: 2016.http://www.libelium.com/development/waspmote/documentation/gasesboard-technical-guide/.Accessed: 2016- 02- 07.

46. Janakiram, D., Reddy, V.A. and Kumar, A.P. (2006). “Outlier detection in wireless sensor networks using Bayesian belief networks”. *First International Conference on Communication System Software and Middleware (Comsware)*. pp. 1-6..

47. Jeffery, S.; Alonso, G.; Franklin, M.; Hong, W.; Widom, J. (2006). “Declarative Support for Sensor Data Cleaning”. *Pervasive Computing*, , .*Springer*: Berlin/Heidelberg, Germany. pp. 83–100.

48. D. Techopedia, "What is Real-Time Data? - Definition from Techopedia", Techopedia.com, 2017. [Online]. Available: https://www.techopedia.com/definition/31256/real-time-data. [Accessed: 03-Jan- 2017].

49. Elshrkawey, M., Elsherif, S.M. and Wahed, M.E.(2018). “An enhancement approach for reducing the energy consumption in wireless sensor networks”. *Journal of King Saud University-Computer and Information Sciences*, 30(2). pp.259-267.

50. Zhang, P., Wang, S., Guo, K. and Wang, J. (2018). “A secure data collection scheme based on compressive sensing in wireless sensor networks”. *Ad Hoc* Networks, 70, pp.73-84.

51. Hess, A.S. and Hess, J.R.(2018). “Principal component analysis”. *Transfusion*. 58(7), 1580-1582.

52. Liu, Y.J., Tran, T., Postma, G., Buydens, L.M.C. and Jansen, J.(2018). “Estimating the number of components and detecting outliers using Angle Distribution of Loading Subspaces (ADLS) in PCA analysis”. *Analytica chimica acta*, 1020, pp.17-29.

53. Berger, P.D., Maurer, R.E. and Celli, G.B.(2018). “Multiple Linear

Regression”. Experimental Design ,pp. 505-532. *Springer*, Cham.

54. Ghrare, S. E., Ali, M. M., Jumari, K., & Ismail, M. (2009). “An efficient low complexity lossless coding algorithm for medical images”. *American Journal of Applied Sciences*, 6(8), 1502.

55. I. B. Arbi, F. Derbel, and F. Strakosch.(2017). “Forecasting methods to reduce energy consumption in WSN,” Proc. *IEEE Instrum. Meas. Technol. Conf.*, pp. 1–6

56. Tayeh, G.B., Makhoul, A., Laiymani, D. and Demerjian, J. (2018). A distributed real-time data prediction and adaptive sensing approach for wireless sensor networks. *Pervasive and Mobile Computing*, 49, pp.62-75.

57. Aderohunmu, Femi A., Giacomo Paci, Davide Brunelli, Jeremiah D. Deng, and Luca Benini.(2013). "Prolonging the lifetime of wireless sensor networks using light-weight forecasting algorithms." *IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, , pp. 461-466.

58. Mesin, L., Aram, S., & Pasero, E. (2014). “A neural data-driven approach to increase Wireless Sensor Networks' lifetime”. *In Computer Applications & Research (WSCAR)*, (pp. 1-3).

59. M. Arunraja and V. Malathi. (2015). “Collective prediction exploiting spatio temporal correlation (CoPeST) for energy efficient wireless sensor networks,” *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 9, no. 7, pp. 2488–2511.

60. Tan, Liansheng, and Mou Wu.(2016). “Data reduction in wireless sensor networks: A hierarchical LMS prediction approach”. *IEEE Sensors Journal* 16, no. 6.pp. 1708-1715.

61. Dias, G. M., Bellalta, B., & Oechsner, S. (2016). “On the importance and feasibility of forecasting data in sensors”. *arXiv preprint arXiv*:1604.01275.

62. F. Strakosch and F. Derbel. (2013).“Fast and efficient dual-forecasting algorithm for wireless sensor networks,” *in Proc. Sensor*, pp 859–863.

63. Fathy, Y., Barnaghi, P., & Tafazolli, R. (2018). “An adaptive method for data reduction in the Internet of Things”. *IEEE 4th World Forum on Internet of Things (WF-IoT)*. pp. 729-735.

64. Harb, H. and Makhoul, A.(2018). Energy-Efficient Sensor Data Collection Approach for Industrial Process Monitoring. *IEEE Transactions on Industrial*

Informatics, 14(2), pp.661-672.

65. Matos, T.B., Brayner, A. and Maia, J.E.B. (2010). "Towards in-network data prediction in wireless sensor networks", *Proceedings of the ACM Symposium on Applied Computing, Sierre, Switzerland*, p.592.

66. Carvalho, C., Gomes, D.G., Agoulmine, N. and de Souza, J.N. (2011) "Improving prediction accuracy for WSN data reduction by applying multivariate spatio-temporal correlation", *Sensors*, Vol. 11, No. 11, pp.10010–10037.

67. Rassam, M. A., Zainal, A., & Maarof, M. A. (2015). "Principal component analysis–based data reduction model for wireless sensor networks". *International Journal of Ad Hoc and Ubiquitous Computing*, 18(1-2), 85-101.

68. Rassam, M. A., Zainal, A., & Maarof, M. A. (2013). "An adaptive and efficient dimension reduction model for multivariate wireless sensor networks applications". *Applied Soft Computing*, 13(4), 1978-1996.

69. Weng, J., Zhang, Y., & Hwang, W. S. (2003). "Candid covariance-free incremental principal component analysis". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8), 1034-1040.

70. Yin, Y., Liu, F., Zhou, X., & Li, Q. (2015). "An efficient data compression model based on spatial clustering and principal component analysis in wireless sensor networks". *Sensors*, 15(8), 19443-19465.

71. Wu, M., Tan, L., & Xiong, N. (2016). "Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications". *Information Sciences*, 329, 800-818.

72. Le Borgne, Y., & Bontempi, G. (2007). "Unsupervised and supervised compression with principal component analysis in wireless sensor networks". *Proceedings of the Workshop on Knowledge Discovery from Data, 13th ACM International Conference on Knowledge Discovery and Data Mining* .pp. 94-103.

73. Le Borgne, Y. A., Raybaud, S., & Bontempi, G. (2008). "Distributed principal component analysis for wireless sensor networks". *Sensors*, 8(8), 4821-4850.

74. Aquino, A. L. L. D. (2011). "A framework for sensor stream reduction in wireless sensor networks".*In The Fifth International Conference on Sensor Technologies and Applications*, pp. 30–35,Nice/Saint Laurent du Var, France.

75. Chen, F., Wen, F., & Jia, H. (2010). "Algorithm of data compression based on multiple principal component analysis over the wsn". *6th International Conference In Wireless Communications Networking and Mobile Computing (WiCOM)*, (pp. 1-4)..

76. Enam, R. N., & Qureshi, R. (2014). "An adaptive data aggregation technique for dynamic cluster based wireless sensor networks". *23rd International Conference In Computer Communication and Networks (ICCCN)*, (pp. 1-7).

77. Alduais, N. A. M., Abdullah, J., & Jamil, A. (2017). "Enhanced Payload Data Reduction Approach for Cluster Head (CH) Nodes". *Telkomnika*, 15(3).

78. Kolo, J. G., Shanmugam, S. A., Lim, D. W. G., & Ang, L. M. (2015). "Fast and efficient lossless adaptive compression scheme for wireless sensor networks". *Computers & Electrical Engineering*, 41, 275-287.

79. Medeiros, H. P., Maciel, M. C., Demo Souza, R., & Pellenz, M. E. (2014). "Lightweight data compression in wireless sensor networks using Huffman coding". *International Journal of Distributed Sensor Networks*, 10(1), 672921.

80. Vecchio, M., Giaffreda, R., & Marcelloni, F. (2014). "Adaptive lossless entropy compressors for tiny IoT devices". *IEEE Transactions on Wireless Communications*, 13(2), 1088-1100.

81. Capo-Chichi, E. P., Guyennet, H., & Friedt, J. M. (2009). "K-RLE: a new data compression algorithm for wireless sensor network". *Third International Conference In Sensor Technologies and Applications (SENSORCOMM'09)*. .pp. 502-507.

82. Marcelloni, F., & Vecchio, M. (2009). "An efficient lossless compression algorithm for tiny nodes of monitoring wireless sensor networks". *The Computer Journal*, 52(8), 969-987.

83. Kasirajan, P., Larsen, C., & Jagannathan, S. (2012). "A new data aggregation scheme via adaptive compression for wireless sensor networks". *ACM Transactions on Sensor Networks (TOSN)*, 9(1), 5.

84. Alsheikh, M. A., Lin, S., Niyato, D., & Tan, H. P. (2016). "Rate-distortion balanced data compression for wireless sensor networks". *IEEE Sensors Journal*, 16(12), 5072-5083.

85. Kim, S., Cho, C., Park, K. J., & Lim, H. (2017). "Increasing network lifetime using data compression in wireless sensor networks with energy harvesting".

International Journal of Distributed Sensor Networks, 13(1), 1550147716689682.

86. Ciuonzo, D., De Maio, A., & Rossi, P. S. (2015). "A systematic framework for composite hypothesis testing of independent Bernoulli trials*". IEEE Signal Processing Letters*, 22(9), 1249-1253.

87. Salvo Rossi, P., Ciuonzo, D., Kansanen, K., & Ekman, T. (2016). "Performance Analysis of Energy Detection for MIMO Decision Fusion in Wireless Sensor Networks Over Arbitrary Fading Channels". *IEEE Transactions on Wireless Communications*, 15(11):7794-7806.

88. Rossi, P. S., Ciuonzo, D., Kansanen, K., & Ekman, T. (2015). "On energy detection for MIMO decision fusion in wireless sensor networks over NLOS fading". *IEEE Communications Letters*, 19(2):303-306.

89. Rafique, Z., Seet, B.C and Al-Anbuky, A. (2013). "Performance analysis of cooperative virtual MIMO systems for wireless sensor networks", *Sensors*, Vol. 13, No. 6, pp.7033–7052.

90. Bartolucci, A. A., Singh, K. P., & Bae, S. (2016). "Descriptive Statistics. Introduction to Statistical Analysis of Laboratory Data*". Hoboken, NJ: John Wiley & Sons.*

91. Iglewicz B, Hoaglin DC. How to Detect and Handle Outliers, ASQC basic references in quality control. Milwaukee, *WI: American Society for Quality Control*. 1993.

92. Contchart.com. (2016). Outlier tests. [online] Available at: https://contchart.com/outliers.aspx [Accessed 21 Mar. 2016].

93. Grubbs, F. E. (1950). "Sample criteria for testing outlying observations." *The Annals of Mathematical Statistic*s, pp.27-58.

94. Tietjen, G.L., Moore, R.H. and Beckman, R.J. (1973). "Testing for a single outlier in simple linear regression." *Technometrics*, 15(4), pp.717-721.

95. Dixon, W.J. (1950). Analysis of extreme values: *Annals of Mathematical Statistics*, 21(4): 488–506.

96. Rosner, B. (1983). "Percentage Points for a Generalized ESD Many-Outlier Procedure." *Technometrics* . 25: pp.165-172.

97. McGill, Robert; Tukey, John W.; Larsen, Wayne A. (1978). "Variations of Box Plots". *The American Statistician*. 32 (1): 12–16. doi:10.2307/2683468.

98. Walsh, J.E. (1959)." Large sample nonparametric rejection of outlying

observations". *Annals of the Institute of Statistical Mathematics*, 10(3), pp.223-232.

99. Zhang, Y., Meratnia, N., & Havinga, P. (2009). "Adaptive and online one-class support vector machine-based outlier detection techniques for wireless sensor networks".*International Conference on Advanced Information Networking and Applications Workshops (WAINA'09).* pp. 990-995.

100. Ghorbel, O., Ayedi, W., Snoussi, H., & Abid, M. (2015). "Fast and efficient outlier detection method in wireless sensor networks*". IEEE sensors journal*, 15(6), 3403-3411.

101. Garcia-Font, V., Garrigues, C., & Rifà-Pous, H. (2016). "A comparative study of anomaly detection techniques for smart city wireless sensor networks". *Sensors*, 16(6), 868.

102. Rassam, M. A., Maarof, M. A., & Zainal, A. (2014). "Adaptive and online data anomaly detection for wireless sensor systems". *Knowledge-Based Systems*, 60, 44-57.

103. Hu, Y., Chen, H., Li, G., Li, H., Xu, R., & Li, J. (2016). "A statistical training data cleaning strategy for the PCA-based chiller sensor fault detection, diagnosis and data reconstruction method". *Energy and Buildings*, 112, 270-278.

104. S. Sivanantham and V. Mohanraj. (2016). "Review Framework for Hybrid Approaches to Outlier Detection in Wireless Sensor Networks" *Middle-East Journal of Scientific Research* 24 (4): 1025-1032.

105. Mahmood, A., Shi, K., Khatoon, S., & Xiao, M. (2013). "Data mining techniques for wireless sensor networks: A survey". *International Journal of Distributed Sensor Networks*, 9(7):, 406316.

106. Ravichandran, J., & Arulappan, A. I. (2013). "Data validation algorithm for wireless sensor networks". *International Journal of Distributed Sensor Networks*, 9(12), 634278.

107. Wander, A. S., Gura, N., Eberle, H., Gupta, V., & Shantz, S. C.(2005). "Energy analysis of public-key cryptography for wireless sensor networks". *Third IEEE International Conference on Pervasive Computing and Communications.(PerCom 2005)*.pp. 324-328

108. Liu, J., Zhang, L., & Sun, R. (2016). "1-RAAP: An efficient 1-round anonymous authentication protocol for wireless body area networks".

Sensors, 16(5): 728.

109. Pathan, A.S.K., Ryu, J.H., Haque, M.M. and Hong, C.S. (2007). “Security management in wireless sensor networks with a public key based scheme”. *Asia-Pacific Network Operations and Management Symposium*.pp. 503-506.

110. I.Data, "Intel Lab Data", db.csail.mit.edu, 2016. [Online]. Available: http://db.csail.mit.edu/labdata/labdata.html. [Accessed: 01- Apr- 2016].

111. L. urban , "Lausanne urban canopy experiment", 2007, [online] Available: http://lcav.epfl.ch/Cms/lang/en/ pid/86035. [Accessed: 01- Apr- 2016].

112. GSB (2007) Grand-St-Bernard (GSB) Dataset, [Online] Available: http://sensorscope.epfl.ch/index.php/%0AEnvironmental Data%3E. [Accessed: 01- Apr- 2016].

113. U. Repository, "UCI Machine Learning Repository: Air Quality Data Set". Archive.ics.uci.edu. N.p., 2016. Web. 4 Aug. 2016.

114. Maureira, M.A., Oldenhof, D. and Teernstra, L. (2015). ‘ThingSpeak – an API and web service for the internet of things’, World WideWeb, Available online:https://staas.home.xs4all.nl/t/swtr/documents/wt2014_thingspeak.pdf (Accessed 29 March, 2018).

115. Halgamuge, M. N., Zukerman, M., Ramamohanarao, K., & Vu, H. L. (2009). “An estimation of sensor energy consumption.” *Progress in Electromagnetics Research*, 12, 259-295.

116. Arunraja, M., Malathi, V., & Sakthivel, E. (2015). “Energy conservation in WSN through multilevel data reduction scheme”. *Microprocessors and Microsystems*, 39(6): 348-357.

117. Miyazaki, T., Li, P., Guo, S., Kitamichi, J., Hayashi, T., & Tsukahara, T. (2015). “On-demand customizable wireless sensor network”. *Procedia Computer Science*, 52, 302-309.

118. Banerjee, A., Mitra, A., Najjar, W. A., Zeinalipour-Yazti, D., Kalogeraki, V., & Gunopulos, D. (2005). “*RISE-Co-S: high performance sensor storage and Co-processing architecture”. SECON*.pp. 1-12.

119 .Ias.ac.in.(2017).Availableat: https://www.ias.ac.in/article/fulltext/reso/021/01/0011-0030 [Accessed 25 Oct. 2017].

120. Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). “An application-specific protocol architecture for wireless microsensor

networks”. *IEEE Transactions on wireless communications*, 1(4): 660-670.

121. Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000). “Energy-efficient communication protocol for wireless microsensor networks”. *Proceedings of the 33rd annual Hawaii international conference on system sciences* .pp1- 10.